Supplemental Document

Biomedical Optics EXPRESS

Deep learning 2D and 3D optical sectioning microscopy using cross-modality Pix2Pix cGAN image translation: supplement

HUIMIN ZHUGE,[1] BRIAN SUMMA,[2] JIHUN HAMM,[2] AND J. QUINCY BROWN[1,*]

[1] *Department of Biomedical Engineering, Tulane University, 500 Lindy Boggs Center, New Orleans, LA 70118, USA*
[2] *Department of Computer Science, Tulane University, New Orleans, LA 70118, USA*
[*] *jqbrown@tulane.edu*

Supplement DOI: https://doi.org/10.6084/m9.figshare.16955512

Parent Article DOI: https://doi.org/10.1364/BOE.439894

Deep learning 2D and 3D optical sectioning microscopy using cross-modality Pix2Pix GAN image translation: supplemental document

This document contains supplemental methods and results.

1. SUPPLEMENTAL METHODS

A. Custom structured illumination microscopy system

Fig. S1 contains a schematic of the custom structured illumination microscopy system used in this work.

B. Evaluation metrics

We chose mean square error (MSE), structural similarity index measurement (SSIM), mutual information (MI), and peak signal-to-noise ratio (PSNR) to quantitatively evaluate the performance of the models between different modalities. The mean square error (MSE) measures the average squared difference between the two images, defined as

$$MSE(X,Y) = \frac{1}{RC}\sum_{i=1}^{R}\sum_{j=1}^{C}(X(i,j) - Y(i,j))^2 \tag{S1}$$

where X and Y denote two images, and R and C denote the row and column of the image, respectively. Structural similarity index (SSIM) measures the similarity between two images based on the degradation of structural information [1].

$$SSIM(X,Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{({\mu_X}^2 + {\mu_Y}^2 + c_1)({\sigma_X}^2 + {\sigma_Y}^2 + c_2)} \\ c_1 = (k_1 L)^2, c_2 = (k_2 L)^2 \tag{S2}$$

where also X and Y denote two images, μ represents the average of the image, σ represents the variance of the image, σ_{XY} represents the covariance of X and Y, L is the dynamic range of the pixel values (typically $2^{NumberBitsPerPixel} - 1$), $k_1 = 0.01$ and $k_2 = 0.03$ by default. The range of SSIM is [0,1], where a higher index means higher similarity between X and Y. Mutual information calculates the mutual dependence between the two images, and can be thought of as the uncertainty reduction about one image given the knowledge of another. Higher MI indicates a large reduction in uncertainty, whereas lower MI indicates a small reduction in uncertainty, and zero MI indicates the two images are independent [2]. MI is given by the equation,

$$MI(X,Y) = \sum_{x,y} P_{XY}(x,y) log \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)} \tag{S3}$$

where $P_{XY}(x,y)$ denotes the joint distribution of X and Y images, $P_X(x)$ and $P_Y(y)$ denotes the marginal distributions of X and Y images, respectively. The PSNR also defined via the MSE, which computes the ratio between the maximum possible pixel value of a image, and the power of the 'corrupted' noise, a higher PSNR refers to a higher similarity between X and Y. It defines by the equation,

$$PSNR(X,Y) = 10\log_{10}\frac{(MAX_X)^2}{MSE(X,Y)} \tag{S4}$$

where MAX_X denotes the maximum possible pixel value of image X. In that case PSNR is expressed as logarithmic quantity with decibel scale.

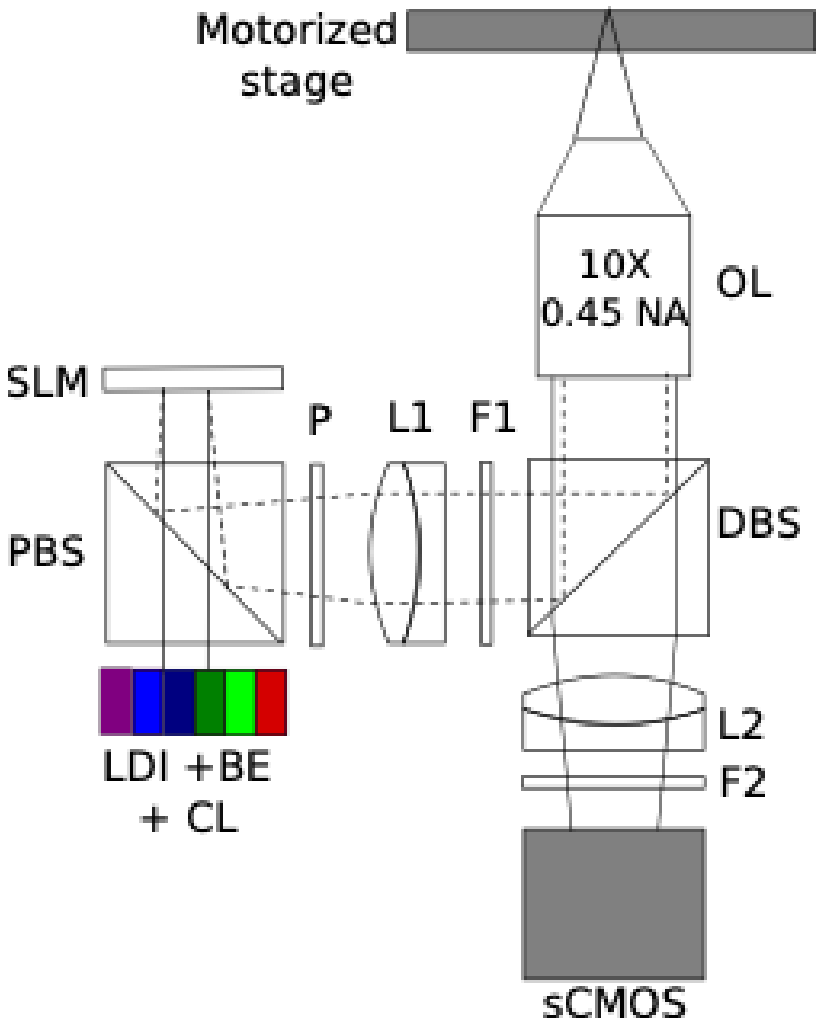

Fig. S1. Schematic of the custom structured illumination system used in this work. SLM: Spatial light modulator. OL: objective lens. DBS: Dichroic beam splitter. PBS: Polarizing beam splitter. L1, L2: tube lenses. F1, F2: excitation and emission filters. LDI: laser illumination engine. BE: beam expander. CL: collimating lens.

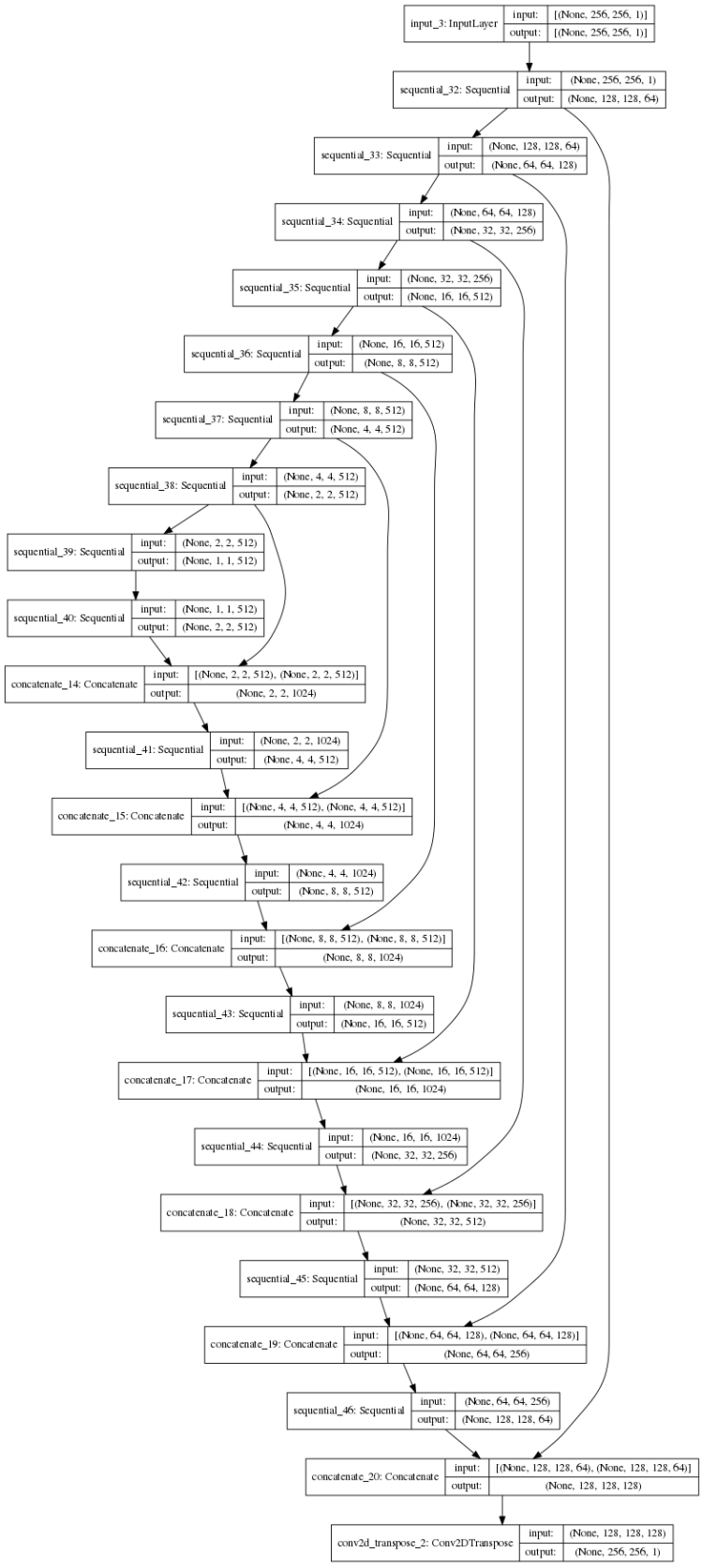

Fig. S2. Model parameters of Generator(U-Net) used in the work.

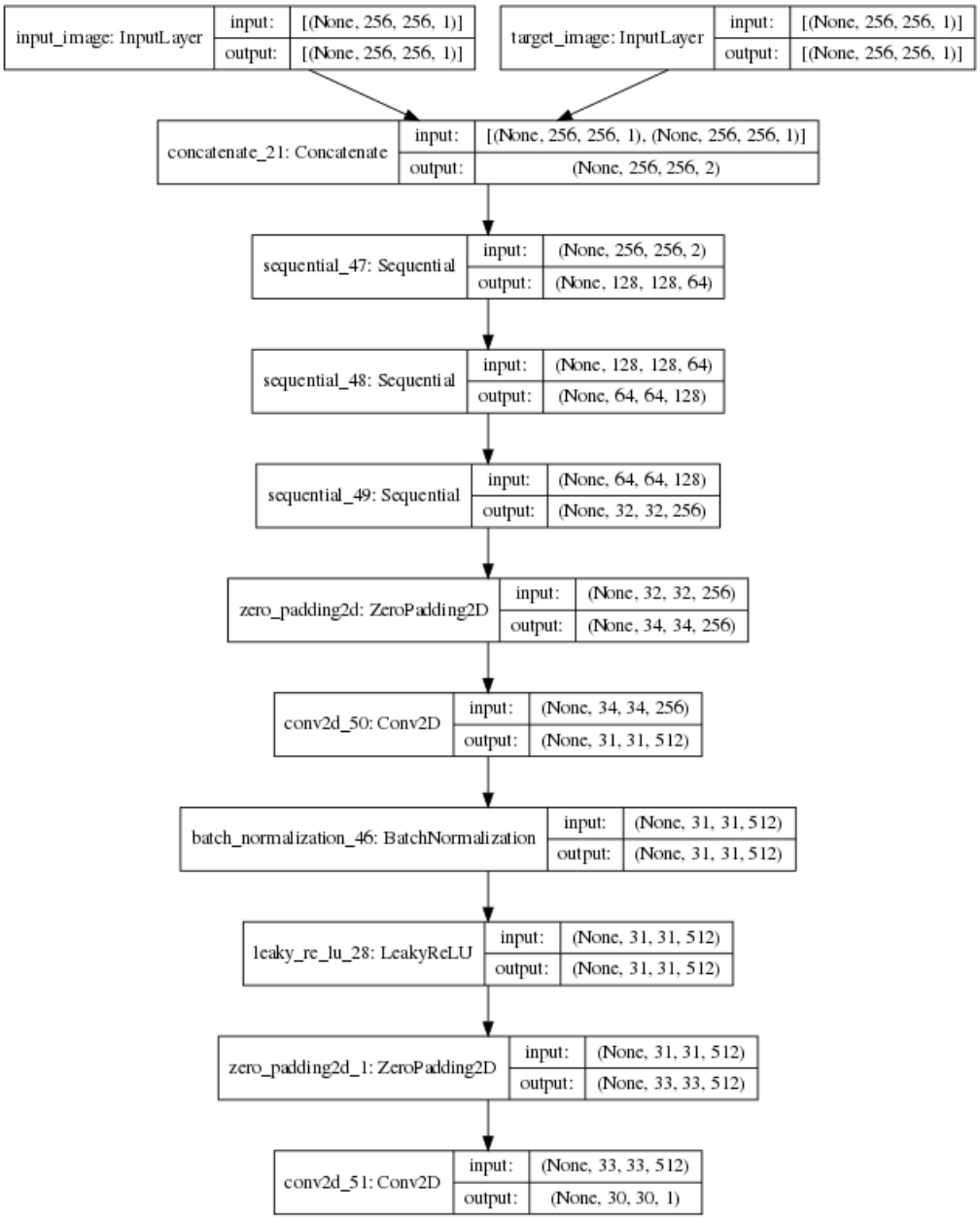

Fig. S3. Model parameters of Discriminator used in the work.

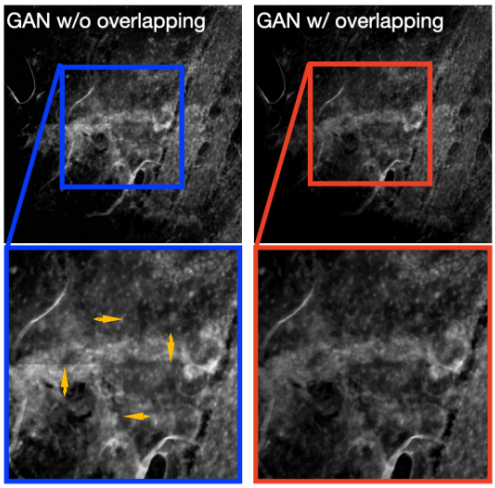

Fig. S4. The use of overlapping patches can enable seamless stitching of predicted images into artifact-free mosaic images. Left: stitched GAN predicted image without overlapping; Right: stitched GAN predicted image with overlapping

C. Model architecture details

Fig. S2 and Fig. S3 shows the detailed parameters of the Generator and Discriminator of the GAN model used in this work. For the CNN model which is constructed by the U-Net only, we preserved the same parameters as Fig. S2.

D. 2D Mosaicking

For many practical applications of optical sectioning microscopy, it is useful to take multiple adjacent images and to stitch them together into a larger mosaic to overcome computational or memory limitations. The most obvious approach is to divide the input images into small, non-overlapping patches. Smaller images will decrease the complexity of the model, as well as the number of parameters. However, computing each image independently and then naively assembling them back together results in obvious artifacts at the boundaries between patches. Fig. S4 illustrates a 2048*2048 image that was generated from four 1024*1024 input patches. The cross-patch artifacts are highlighted in yellow arrows. We therefore designed a pipeline to ensure that overall mosaic image quality will be preserved when using smaller patches, shown in Fig. S5. First, we split the wide-field input image into small patches of any size (1024*1024 for this instance) with 50% overlapping. For example, for each 1024*1024 patch (center white dashed square), we also retrieved the four 1024*1024 neighbor patches with 50% overlapping (red, blue, green, yellow dashed squares). Second, each patch was passed through the previously-trained GAN model to obtain the corresponding predicted image. In this example, all five patches went through the well-trained model, to get the corresponding predicted patches. Third, we stitched the patches back together by selecting the relative center region of each patch. For this particular example, we preserved center of white square [256:768,256:768], center right of red square [256:1024,512:768], center bottom of blue square [512:768,1:768], center left of green square [1:768,256:512], center top of yellow square [256:512,256:1024], then stitched the five small patches back according to the original coordinates.

The right of Fig. S4 shows the result performed with the overlapping procedure. Compared with the left image without the overlapping, we do not see obvious lines dividing the images, and the edge effects are removed. In this way, not only are the edge effects removed for mosaic reconstruction, but the sub-patching procedure of the pipeline enables reconstruction of any large scale mosaic image even with non-optimal hardware.

2. SUPPLEMENTAL RESULTS

A. Application to images from different sample types

To test the universality of the approach to different sample types, we applied the model trained on images of the surface of the prostate (prostatectomy specimen) to images collected from prostate

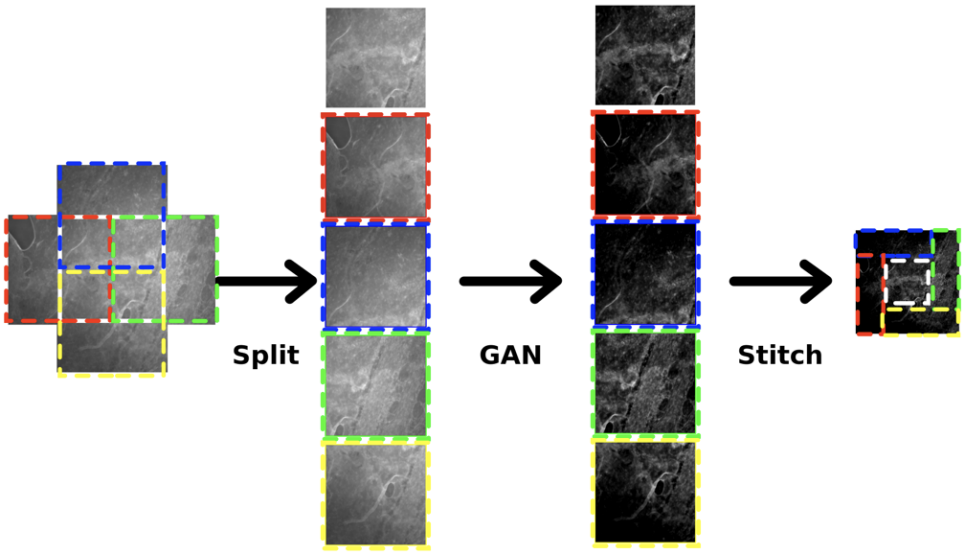

Fig. S5. Overlapping patch and stitching pipeline to reduce edge effects in final mosaic images. Edge colored rectangles are consistent through the pipeline.

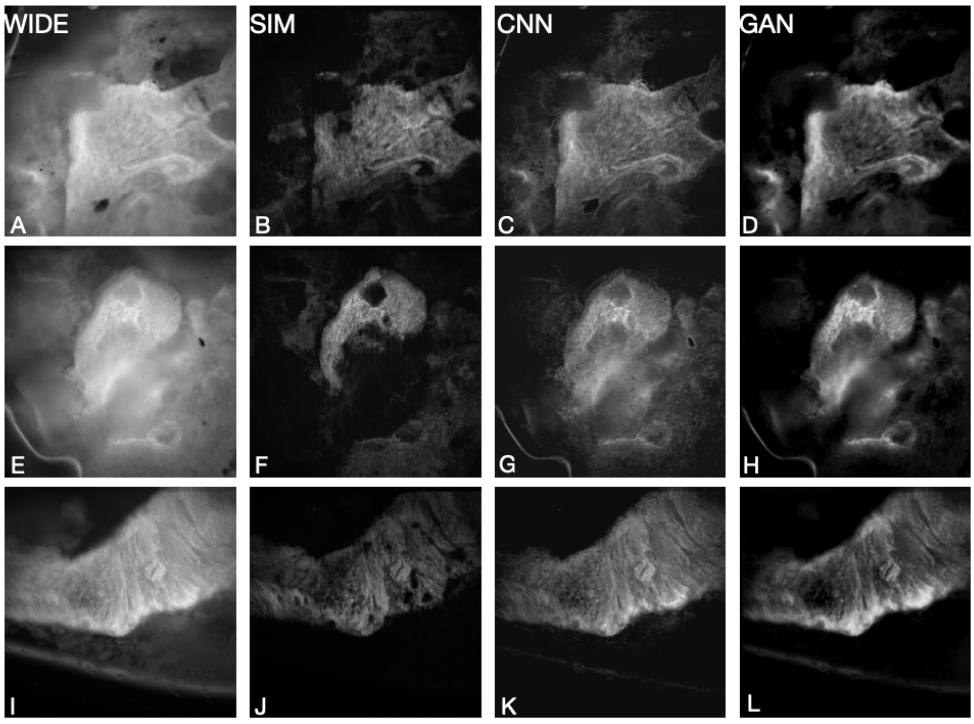

Fig. S6. Application of the CNN and GAN models to a prostate needle-core biopsy (prostate parenchyma) using a different fluorescent stain. From left to right: wide-field input images (WIDE, A, E, I), SIM ground truth images (SIM, B, E, J), , CNN predicted images (CNN, C, G, K), and GAN predicted images (GAN, D, H, L).

biopsies, which are collected from inside the prostate, not the surface. Even though both types of samples are from prostate, it is known that the microarchitecture of the surface of the prostate is very different from the internal prostate parenchyma. In addition, these samples were stained with a different fluorescent stain (eosin vs. acridine orange), which stain the tissue microarchitecture differently. Our results in Fig. S6 show that even when applied to test images that are different from the training images in terms of tissue structure and contrast mechanism, the GAN model is qualitatively superior to the CNN model in terms of optical sectioning performance, compared to the ground truth SIM images. However, these results were less accurate than for the case where the model is used to predict images from the similar type (tissue and stain technique) present in the training set. This confirms a limitation of supervised deep learning models using paired training datasets, which however could potentially be alleviated by training the model with images from a wide variety of tissue types and staining protocols.

REFERENCES

1. Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Process. **13**, 600–612 (2004).
2. D. B. Russakoff, C. Tomasi, T. Rohlfing, and C. R. Maurer, "Image similarity using mutual information of regions," in *Computer Vision - ECCV 2004*, T. Pajdla and J. Matas, eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2004), pp. 596–607.